



IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/825,198

DATE: 08/04/2004

TIME: 08:31:42

Input Set : N:\Crif3\RULE60\10825198.raw

Output Set: N:\CRF4\08042004\J825198.raw

1 <110> APPLICANT: Glucksman, Maria Alexandra  
 2 Williamson, Mark  
 3 Tsia, Fong-Ying  
 4 Rudolph-Owen, Laura A.  
 5 <120> TITLE OF INVENTION: 22438, 23553, 25278, and 26212 Novel  
 6 Human Sulfatases (A CIP Application)  
 7 <130> FILE REFERENCE: 35800/208398(5800-79  
 8 <140> CURRENT APPLICATION NUMBER: US/10/825,198  
 9 <141> CURRENT FILING DATE: 2004-04-15  
 10 <150> PRIOR APPLICATION NUMBER: US/10/314,881  
 11 <151> PRIOR FILING DATE: 2002-12-09  
 12 <150> PRIOR APPLICATION NUMBER: US/09/773,426  
 13 <151> PRIOR FILING DATE: 2001-01-31  
 14 <150> PRIOR APPLICATION NUMBER: US 09/495,823  
 15 <151> PRIOR FILING DATE: 2000-01-31  
 16 <160> NUMBER OF SEQ ID NOS: 14  
 17 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 19 <210> SEQ ID NO: 1  
 20 <211> LENGTH: 525  
 21 <212> TYPE: PRT  
 22 <213> ORGANISM: homo sapiens  
 23 <400> SEQUENCE: 1  
 24 Met Gly Trp Leu Phe Leu Lys Val Leu Leu Ala Gly Val Ser Phe Ser  
 25 1 5 10 15  
 26 Gly Phe Leu Tyr Pro Leu Val Asp Phe Cys Ile Ser Gly Lys Thr Arg  
 27 20 25 30  
 28 Gly Gln Lys Pro Asn Phe Val Ile Ile Leu Ala Asp Asp Met Gly Trp  
 29 35 40 45  
 30 Gly Asp Leu Gly Ala Asn Trp Ala Glu Thr Lys Asp Thr Ala Asn Leu  
 31 50 55 60  
 32 Asp Lys Met Ala Ser Glu Gly Met Arg Phe Val Asp Phe His Ala Ala  
 33 65 70 75 80  
 34 Ala Ser Thr Cys Ser Pro Ser Arg Ala Ser Leu Leu Thr Gly Arg Leu  
 35 85 90 95  
 36 Gly Leu Arg Asn Gly Val Thr Arg Asn Phe Ala Val Thr Ser Val Gly  
 37 100 105 110  
 38 Gly Leu Pro Leu Asn Glu Thr Thr Leu Ala Glu Val Leu Gln Gln Ala  
 39 115 120 125  
 40 Gly Tyr Val Thr Gly Ile Ile Gly Lys Trp His Leu Gly His His Gly  
 41 130 135 140  
 42 Ser Tyr His Pro Asn Phe Arg Gly Phe Asp Tyr Tyr Phe Gly Ile Pro  
 43 145 150 155 160  
 44 Tyr Ser His Asp Met Gly Cys Thr Asp Thr Pro Gly Tyr Asn His Pro

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/825,198

DATE: 08/04/2004

TIME: 08:31:42

Input Set : N:\Crif3\RULE60\10825198.raw

Output Set: N:\CRF4\08042004\J825198.raw

```

45          165          170          175
46  Pro Cys Pro Ala Cys Pro Gln Gly Asp Gly Pro Ser Arg Asn Leu Gln
47          180          185          190
48  Arg Asp Cys Tyr Thr Asp Val Ala Leu Pro Leu Tyr Glu Asn Leu Asn
49          195          200          205
50  Ile Val Glu Gln Pro Val Asn Leu Ser Ser Leu Ala Gln Lys Tyr Ala
51          210          215          220
52  Glu Lys Ala Thr Gln Phe Ile Gln Arg Ala Ser Thr Ser Gly Arg Pro
53  225          230          235          240
54  Phe Leu Leu Tyr Val Ala Leu Ala His Met His Val Pro Leu Pro Val
55          245          250          255
56  Thr Gln Leu Pro Ala Ala Pro Arg Gly Arg Ser Leu Tyr Gly Ala Gly
57          260          265          270
58  Leu Trp Glu Met Asp Ser Leu Val Gly Gln Ile Lys Asp Lys Val Asp
59          275          280          285
60  His Thr Val Lys Glu Asn Thr Phe Leu Trp Phe Thr Gly Asp Asn Gly
61          290          295          300
62  Pro Trp Ala Gln Lys Cys Glu Leu Ala Gly Ser Val Gly Pro Phe Thr
63  305          310          315          320
64  Gly Phe Trp Gln Thr Arg Gln Gly Gly Ser Pro Ala Lys Gln Thr Thr
65          325          330          335
66  Trp Glu Gly Gly His Arg Val Pro Ala Leu Ala Tyr Trp Pro Gly Arg
67          340          345          350
68  Val Pro Val Asn Val Thr Ser Thr Ala Leu Leu Ser Val Leu Asp Ile
69          355          360          365
70  Phe Pro Thr Val Val Ala Leu Ala Gln Ala Ser Leu Pro Gln Gly Arg
71          370          375          380
72  Arg Phe Asp Gly Val Asp Val Ser Glu Val Leu Phe Gly Arg Ser Gln
73  385          390          395          400
74  Pro Gly His Arg Val Leu Phe His Pro Asn Ser Gly Ala Ala Gly Glu
75          405          410          415
76  Phe Gly Ala Leu Gln Thr Val Arg Leu Glu Arg Tyr Lys Ala Phe Tyr
77          420          425          430
78  Ile Thr Gly Gly Ala Arg Ala Cys Asp Gly Ser Thr Gly Pro Glu Leu
79          435          440          445
80  Gln His Lys Phe Pro Leu Ile Phe Asn Leu Glu Asp Asp Thr Ala Glu
81          450          455          460
82  Ala Val Pro Leu Glu Arg Gly Gly Ala Glu Tyr Gln Ala Val Leu Pro
83  465          470          475          480
84  Glu Val Arg Lys Val Leu Ala Asp Val Leu Gln Asp Ile Ala Asn Asp
85          485          490          495
86  Asn Ile Ser Ser Ala Asp Tyr Thr Gln Asp Pro Ser Val Thr Pro Cys
87          500          505          510
88  Cys Asn Pro Tyr Gln Ile Ala Cys Arg Cys Gln Ala Ala
89          515          520          525
91 <210> SEQ ID NO: 2
92 <211> LENGTH: 2175
93 <212> TYPE: DNA
94 <213> ORGANISM: homo sapiens

```

## RAW SEQUENCE LISTING

DATE: 08/04/2004

PATENT APPLICATION: US/10/825,198

TIME: 08:31:42

Input Set : N:\Crif3\RULE60\10825198.raw

Output Set: N:\CRF4\08042004\J825198.raw

```

95 <220> FEATURE:
96 <221> NAME/KEY: CDS
97 <222> LOCATION: (248)...(1825)
98 <400> SEQUENCE: 2
99   cacgcgtccg caaatttctt gattcttttg aattaggatt ccagatgggg gcctcatttc      60
100   tacagccccc aacatttcta tagccgttat cactgccatc accactgccca ccagcatctt      120
101   cttgcagatt ccaccctgcg tcccagaga cttcctgctt tgaaagtgag cagaaaggaa      180
102   gctctcagaa aaatctctag tgggtggctgc cgtcgctcca gacaatcgga atcctgcctt      240
103   caccacc atg ggc tgg ctt ttt cta aag gtt ttg ttg gcg gga gtg agt      289
104   Met Gly Trp Leu Phe Leu Lys Val Leu Leu Ala Gly Val Ser
105   1          5          10
106   ttc tca gga ttt ctt tat cct ctt gtg gat ttt tgc atc agt ggg aaa      337
107   Phe Ser Gly Phe Leu Tyr Pro Leu Val Asp Phe Cys Ile Ser Gly Lys
108   15          20          25          30
109   aca aga gga cag aag cca aac ttt gtg att att ttg gcc gat gac atg      385
110   Thr Arg Gly Gln Lys Pro Asn Phe Val Ile Ile Leu Ala Asp Asp Met
111   35          40          45
112   ggg tgg ggt gac ctg gga gca aac tgg gca gaa aca aag gac act gcc      433
113   Gly Trp Gly Asp Leu Gly Ala Asn Trp Ala Glu Thr Lys Asp Thr Ala
114   50          55          60
115   aac ctt gat aag atg gct tcg gag gga atg agg ttt gtg gat ttc cat      481
116   Asn Leu Asp Lys Met Ala Ser Glu Gly Met Arg Phe Val Asp Phe His
117   65          70          75
118   gca gct gcc tcc acc tgc tca ccc tcc cgg gct tcc ttg ctc acc ggc      529
119   Ala Ala Ala Ser Thr Cys Ser Pro Ser Arg Ala Ser Leu Leu Thr Gly
120   80          85          90
121   cgg ctt ggc ctt cgc aat gga gtc aca cgc aac ttt gca gtc act tct      577
122   Arg Leu Gly Leu Arg Asn Gly Val Thr Arg Asn Phe Ala Val Thr Ser
123   95          100          105          110
124   gtg gga ggc ctt ccg ctc aac gag acc acc ttg gca gag gtg ctg cag      625
125   Val Gly Gly Leu Pro Leu Asn Glu Thr Thr Leu Ala Glu Val Leu Gln
126   115          120          125
127   cag gcg ggt tac gtc act ggg ata ata ggc aaa tgg cat ctt gga cac      673
128   Gln Ala Gly Tyr Val Thr Gly Ile Ile Gly Lys Trp His Leu Gly His
129   130          135          140
130   cac ggc tct tat cac ccc aac ttc cgt ggt ttt gat tac tac ttt gga      721
131   His Gly Ser Tyr His Pro Asn Phe Arg Gly Phe Asp Tyr Tyr Phe Gly
132   145          150          155
133   atc cca tat agc cat gat atg ggc tgt act gat act cca ggc tac aac      769
134   Ile Pro Tyr Ser His Asp Met Gly Cys Thr Asp Thr Pro Gly Tyr Asn
135   160          165          170
136   cac cct cct tgt cca gcg tgt cca cag ggt gat gga cca tca agg aac      817
137   His Pro Pro Cys Pro Ala Cys Pro Gln Gly Asp Gly Pro Ser Arg Asn
138   175          180          185          190
139   ctt caa aga gac tgt tac act gac gtg gcc ctc cct ctt tat gaa aac      865
140   Leu Gln Arg Asp Cys Tyr Thr Asp Val Ala Leu Pro Leu Tyr Glu Asn
141   195          200          205
142   ctc aac att gtg gag cag ccg gtg aac ttg agc agc ctt gcc cag aag      913
143   Leu Asn Ile Val Glu Gln Pro Val Asn Leu Ser Ser Leu Ala Gln Lys

```

## RAW SEQUENCE LISTING

DATE: 08/04/2004

PATENT APPLICATION: US/10/825,198

TIME: 08:31:42

Input Set : N:\Crif3\RULE60\10825198.raw

Output Set: N:\CRF4\08042004\J825198.raw

144		210		215		220		
145	tat gct gag aaa gca acc cag ttc atc cag cgt gca agc acc agc ggg							961
146	Tyr Ala Glu Lys Ala Thr Gln Phe Ile Gln Arg Ala Ser Thr Ser Gly							
147		225		230		235		
148	agg ccc ttc ctg ctg tat gtg gct ctg gcc cac atg cac gtg ccc tta							1009
149	Arg Pro Phe Leu Leu Tyr Val Ala Leu Ala His Met His Val Pro Leu							
150		240		245		250		
151	ccc gtg act cag cta cca gca gcg cca cgg ggc aga agc ctg tat ggt							1057
152	Pro Val Thr Gln Leu Pro Ala Ala Pro Arg Gly Arg Ser Leu Tyr Gly							
153		255		260		265		270
154	gca ggg ctg tgg gag atg gac agt ctg gtg ggc cag atc aag gac aaa							1105
155	Ala Gly Leu Trp Glu Met Asp Ser Leu Val Gly Gln Ile Lys Asp Lys							
156		275		280		285		
157	gtt gac cac aca gtg aag gaa aac aca ttc ctg tgg ttt aca gga gac							1153
158	Val Asp His Thr Val Lys Glu Asn Thr Phe Leu Trp Phe Thr Gly Asp							
159		290		295		300		
160	aat ggc ccg tgg gct cag aag tgt gag cta gcg ggc agt gtg ggt ccc							1201
161	Asn Gly Pro Trp Ala Gln Lys Cys Glu Leu Ala Gly Ser Val Gly Pro							
162		305		310		315		
163	ttc act gga ttt tgg caa act cgt caa ggg gga agt cca gcc aag cag							1249
164	Phe Thr Gly Phe Trp Gln Thr Arg Gln Gly Gly Ser Pro Ala Lys Gln							
165		320		325		330		
166	acg acc tgg gaa gga ggg cac cgg gtc cca gca ctg gct tac tgg cct							1297
167	Thr Thr Trp Glu Gly Gly His Arg Val Pro Ala Leu Ala Tyr Trp Pro							
168		335		340		345		350
169	ggc aga gtt cca gtt aat gtc acc agc act gcc ttg tta agc gtg ctg							1345
170	Gly Arg Val Pro Val Asn Val Thr Ser Thr Ala Leu Leu Ser Val Leu							
171		355		360		365		
172	gac att ttt cca act gtg gta gcc ctg gcc cag gcc agc tta cct caa							1393
173	Asp Ile Phe Pro Thr Val Val Ala Leu Ala Gln Ala Ser Leu Pro Gln							
174		370		375		380		
175	gga cgg cgc ttt gat ggt gtg gac gtc tcc gag gtg ctg ttt ggc cgg							1441
176	Gly Arg Arg Phe Asp Gly Val Asp Val Ser Glu Val Leu Phe Gly Arg							
177		385		390		395		
178	tca cag cct ggg cac agg gtg ctg ttc cac ccc aac agc ggg gca gct							1489
179	Ser Gln Pro Gly His Arg Val Leu Phe His Pro Asn Ser Gly Ala Ala							
180		400		405		410		
181	gga gag ttt gga gcc ctg cag act gtc cgc ctg gag cgt tac aag gcc							1537
182	Gly Glu Phe Gly Ala Leu Gln Thr Val Arg Leu Glu Arg Tyr Lys Ala							
183		415		420		425		430
184	ttc tac att acc ggt gga gcc agg gcg tgt gat ggg agc acg ggg cct							1585
185	Phe Tyr Ile Thr Gly Gly Ala Arg Ala Cys Asp Gly Ser Thr Gly Pro							
186		435		440		445		
187	gag ctg cag cat aag ttt cct ctg att ttc aac ctg gaa gac gat acc							1633
188	Glu Leu Gln His Lys Phe Pro Leu Ile Phe Asn Leu Glu Asp Asp Thr							
189		450		455		460		
190	gca gaa gct gtg ccc cta gaa aga ggt ggt gcg gag tac cag gct gtg							1681
191	Ala Glu Ala Val Pro Leu Glu Arg Gly Gly Ala Glu Tyr Gln Ala Val							
192		465		470		475		

## RAW SEQUENCE LISTING

DATE: 08/04/2004

PATENT APPLICATION: US/10/825,198

TIME: 08:31:42

Input Set : N:\Crif3\RULE60\10825198.raw

Output Set: N:\CRF4\08042004\J825198.raw

```

193   ctg ccc gag gtc aga aag gtt ctt gca gac gtc ctc caa gac att gcc      1729
194   Leu Pro Glu Val Arg Lys Val Leu Ala Asp Val Leu Gln Asp Ile Ala
195           480                      485                      490
196   aac gac aac atc tcc agc gca gat tac act cag gac cct tca gta act      1777
197   Asn Asp Asn Ile Ser Ser Ala Asp Tyr Thr Gln Asp Pro Ser Val Thr
198   495                      500                      505                      510
199   ccc tgc tgt aat ccc tac caa att gcc tgc cgc tgt caa gcc gca taa      1825
200   Pro Cys Cys Asn Pro Tyr Gln Ile Ala Cys Arg Cys Gln Ala Ala *
201           515                      520                      525
202   cagaccaatt tttattccac gaggaggagt acctggaaat taggcaagtt tgcttccaaa      1885
203   tttcattttt accctcttta caaacacacg ctttagttta gtcttggagt ttagtitttg      1945
204   agttagcctt gcatatccct tctgtatcct gtccctcctc cacgccgacc cgagagcagc      2005
205   tgagctgcgc tggctctggg cagggagtgt gccttaatgg gaagcacacg ggctttggag      2065
206   tcaggcacag gtgccagctc cagcttttga acttgggcaa ttgtttaacc taacctgcaa      2125
207   gttgattttg agggttaaat aaaggcatac atgaaaaaaaa aaaaaaaaaa      2175
209 <210> SEQ ID NO: 3
210 <211> LENGTH: 871
211 <212> TYPE: PRT
212 <213> ORGANISM: homo sapiens
213 <400> SEQUENCE: 3
214   Met Lys Tyr Ser Cys Cys Ala Leu Val Leu Ala Val Leu Gly Thr Glu
215           1           5           10           15
216   Leu Leu Gly Ser Leu Cys Ser Thr Val Arg Ser Pro Arg Phe Arg Gly
217           20           25           30
218   Arg Ile Gln Gln Glu Arg Lys Asn Ile Arg Pro Asn Ile Ile Leu Val
219           35           40           45
220   Leu Thr Asp Asp Gln Asp Val Glu Leu Gly Ser Leu Gln Val Met Asn
221           50           55           60
222   Lys Thr Arg Lys Ile Met Glu His Gly Gly Ala Thr Phe Ile Asn Ala
223           65           70           75           80
224   Phe Val Thr Thr Pro Met Cys Cys Pro Ser Arg Ser Ser Met Leu Thr
225           85           90           95
226   Gly Lys Tyr Val His Asn His Asn Val Tyr Thr Asn Asn Glu Asn Cys
227           100          105          110
228   Ser Ser Pro Ser Trp Gln Ala Met His Glu Pro Arg Thr Phe Ala Val
229           115          120          125
230   Tyr Leu Asn Asn Thr Gly Tyr Arg Thr Ala Phe Phe Gly Lys Tyr Leu
231           130          135          140
232   Asn Glu Tyr Asn Gly Ser Tyr Ile Pro Pro Gly Trp Arg Glu Trp Leu
233           145          150          155          160
234   Gly Leu Ile Lys Asn Ser Arg Phe Tyr Asn Tyr Thr Val Cys Arg Asn
235           165          170          175
236   Gly Ile Lys Glu Lys His Gly Phe Asp Tyr Ala Lys Asp Tyr Phe Thr
237           180          185          190
238   Asp Leu Ile Thr Asn Glu Ser Ile Asn Tyr Phe Lys Met Ser Lys Arg
239           195          200          205
240   Met Tyr Pro His Arg Pro Val Met Met Val Ile Ser His Ala Ala Pro
241           210          215          220
242   His Gly Pro Glu Asp Ser Ala Pro Gln Phe Ser Lys Leu Tyr Pro Asn

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/825,198

DATE: 08/04/2004  
TIME: 08:31:43

Input Set : N:\Crf3\RULE60\10825198.raw  
Output Set: N:\CRF4\08042004\J825198.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; N Pos. 4310

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/825,198

DATE: 08/04/2004

TIME: 08:31:43

Input Set : N:\Crf3\RULE60\10825198.raw

Output Set: N:\CRF4\08042004\J825198.raw

L:529 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:4265